Cancelling of Insurance Contracts

Brekelmans, R.C.M.; De Waegenaere, A.M.B.

Publication date:
1999

Citation for published version (APA):
Brekelmans, R. C. M., & De Waegenaere, A. M. B. (1999). *Cancelling of Insurance Contracts*. (CentER Discussion Paper; Vol. 1999-60). Operations research.

Canceling of Insurance Contracts

RUUD BREKELMANS* ANJA DE WAEGENAERE†

July 2, 1999

Abstract

In this paper we consider an insurer who has incomplete information about the claim frequency of the risk process. He therefore calculates the premium on the basis of a prior distribution for the claim frequency. Future information might then reveal that it is no longer optimal for the insurer to continue to offer the insurance under the current conditions. We consider a model where, at certain points in time, the insurer can decide to cancel the insurance, possibly at the expense of canceling costs. The model is applicable to long-period, client tailored insurance contracts as well as insurance offered to large groups of insureds on a single period basis. We derive the optimal canceling policy and analyze the influence of the different model parameters on the expected lifetime of the insurance, the insurer's expected surplus, and the safety loading.

Keywords: insurance, unknown claim frequency, canceling, stochastic dynamic programming, Bayesian updating.

AMS subject classification: 90A46, 90C39.

*Department of Econometrics, Tilburg University.

†Corresponding author: CentER for Economic Research, and Department of Econometrics, Tilburg University, P.O. Box 90153, 5000 LE Tilburg, The Netherlands. The authors wish to thank Frank van der Duyn Schouten for valuable comments and discussions.

1 Introduction

The classical risk process, as introduced in [3], describes the evolution over time of the surplus generated by an insurance portfolio. This surplus consists of premium income reduced with claim amounts. Claims arrive according to a homogeneous Poisson process, and claim heights are i.i.d. and independent of the claim arrival process. The aggregate premium income over a time interval of length t equals pt, where p denotes the *premium density*. It is determined such that, for any given time interval, the expected surplus equals a *safety loading* θ times the expected aggregate claim amount. For given claim frequency and claim height distribution, the *safety loading* can be varied to affect the stability of the portfolio. Several criteria for stability exist, among which the *probability of ruin* is the most prominent one (see for example [4], for an overview on the literature).

In the above described classical risk process, the insurer offers insurance at a given, constant, premium density, for the whole length of the planning horizon. Stability criteria can then be determined on the basis of this assumption. This assumption is relevant in cases where the claim frequency and the claim height distribution are constant over time, and known to the insurer at the start of the planning horizon. When the above assumptions are not satisfied, stability is also largely influenced by the insurer's policy with regard to maintaining the current premium system. Grandell [7] therefore describes a generalized risk model where claim frequency is time dependent, and the premium density is continuously adjusted to new estimates of the claim frequency.

We consider a situation where claim frequency is constant, but unknown to the insurer. This may in particular be the case for development of new, or client tailored, insurance products, where little historical data is available to the insurer. An estimate of the claim frequency is then used to decide on contract specifications. After some time, observed information might indicate that it is no longer optimal to continue to offer the insurance under the current conditions. However, estimates of the claim frequency can become more accurate over time, and, moreover, the decision to stop offering the insurance can be costly to the insurer. Therefore, in a discrete time framework, a trade-off arises between continuing with the premium system one more period, so that in the next period the decision can be made with more reliable information, and stopping to offer the insurance now, at higher canceling cost.

Our aim is to determine the optimal strategy with regard to canceling. At each decision moment, the optimal decision rule weighs, conditional on the history of the risk process, the expected future surplus that results from continuing the contract—taking into account all available options at future decisions moments—against the canceling costs. With regard to the cost of canceling we consider two situations, depending on contract specifications. In the first case, which arises typically in case of large, client tailored, insurance contracts, the insurer in principal commits to insure the risk for a given number of time periods, but has the option to cancel earlier. Exercising this option however implies that the insured has to be compensated for the fact that he will have to renegotiate insurance for the remaining periods. In the second case, which occurs in case of insurance offered to a large group of insureds on a yearly basis, the insurer has no commitment to offer the insurance for more than one year, and can therefore cancel without any compensation to the insured.

If the option to cancel the contract is available to the insurer, this clearly affects the value of the portfolio for him. In case of precommitment, it also affects the insured's expected utility of the insurance offered to him. Whether for the insurer a premium density is sufficient to cover the risk, and whether for the insured, it is acceptable, will depend on the *effective safety loading*, which equals the fraction of expected total surplus over expected total costs at the end of the contract, given that the optimal canceling strategy is applied. Since total costs consist of claim amounts augmented with canceling costs, and since the option to cancel is available, the effective safety loading in general differs from the safety loading used to determine the premium density.

We proceed as follows. In Section 2 the general model is presented. In Section 3 we use stochastic dynamic programming to derive the optimal canceling strategy for the case where the insurer uses Bayesian updating of the claim frequency distribution. The optimal decision rule is of the following form. For each future decision moment there is a threshold value such that the insurer should cancel iff the past claim frequency exceeds this value. The threshold value at a certain time instance depends on: the remaining number of periods in the planning horizon, the accuracy of the prior distribution, the safety loading, and the cancellation costs.

In Section 4, we determine the *effective safety loading*. Section 5 concludes.

2 General model

We consider the classical risk process, as in e.g. [3] or [5], over planning horizon $[0, T]$. Let $\{N_t : t \geq 0\}$ denote the claim arrival process, i.e. N_t denotes the number of claims that arrive during the period $[0, t]$. The claim arrival process is a homogeneous Poisson process with parameter $\lambda > 0$. The claim amounts $\{X_n : n \in \mathbb{N}\}$ are i.i.d. and independent of the claim arrival process. Finally, S_t denotes the aggregate claim amount up to epoch t, i.e. $\{S_t : t \geq 0\}$ is a compound Poisson process:

$$S_t = \sum_{n=1}^{N_t} X_n.$$

Let $P(t_1, t_2)$ denote the premium received for the risk covered by the insurer in the time interval $(t_1, t_2]$, $(0 \leq t_1 < t_2 \leq T)$. The premium is determined using the expected value premium principle (see e.g. [6]). This implies that the pure premium, i.e. premium net of transaction costs, taxes, etc., is proportional to the expected value of the insured risk—at the time the contract is signed. With given claim frequency λ, this implies that:

$$\begin{aligned} P(t_1, t_2) &= (1 + \theta)E[S_{t_2} - S_{t_1}], \\ &= (1 + \theta)\lambda\mu(t_2 - t_1), \\ &= p(t_2 - t_1), \end{aligned}$$

where $\theta \geq 0$ denotes the *safety loading*, $\mu = E[X_i]$ is the expected claim height, and $p = (1 + \theta)\lambda\mu$ denotes the *premium density* (see e.g. [2]).

The insurer however does not know the exact value of the claim frequency, λ, and therefore uses a prior distribution:

$$\Lambda \sim Gamma(t_0, \lambda_0 t_0),$$

to calculate the premium to cover the risk in $[0, T]$, so that the premium density used at date zero equals:

$$
\begin{aligned}
p &= (1+\theta)E[\Lambda]\mu, \\
&= (1+\theta)\lambda_0\mu.
\end{aligned}
$$

The above premium is determined using the information available at date zero. At time t, new information about the claim arrival process is available. Let $\mathcal{H}_t$ denote the history of the claim arrival process up to time t, i.e.

$$
\mathcal{H}_t = \sigma\{N_u : 0 \leq u \leq t\}.
$$

This information can be used to update the initial prior distribution to a posterior distribution using the theory of Bayesian updating. Given that N_t claims have arrived in $[0,t]$, the posterior distribution is as follows (see e.g. [1]):

$$
\Lambda|\mathcal{H}_t \sim Gamma\left(t_0+t, \lambda_0 t_0 + N_t\right). \tag{1}
$$

In the sequel we will denote $E[\cdot|\mathcal{H}_t]$ for the expectation with respect to the posterior distribution at time t. Based on this posterior distribution, the expected claim frequency equals:

$$
\lambda_t = E[\Lambda|\mathcal{H}_t] = \frac{\lambda_0 t_0 + N_t}{t_0+t}, \tag{2}
$$

so that, for any $u \geq t$, the expected value of aggregate claim amounts in $(t,u]$ equals:

$$
E[S_u - S_t|\mathcal{H}_t] = \lambda_t\mu(u-t).
$$

Now, since premium income is fixed, the estimated claim frequency at time t implies that, for any future time interval $(t,u]$, the estimated safety loading θ_t satisfies:

$$
\begin{aligned}
& p(t-u) = (1+\theta_t)E[S_u - S_t|\mathcal{H}_t], \\
\Leftrightarrow \quad & (1+\theta)\lambda_0\mu = (1+\theta_t)\lambda_t\mu, \\
\Leftrightarrow \quad & \theta_t = (1+\theta)\frac{\lambda_0}{\lambda_t} - 1,
\end{aligned}
$$

so that θ_t is lower than θ whenever λ_t is higher than λ_0. Therefore, depending on the value of λ_t, it may be optimal for the insurer to stop offering the insurance at the given premium density. This decision however will also depend on contract specifications. In case of precommitment, the insurer may need to pay a cancellation fee to compensate for the renegotiation costs the insured will incur in order to obtain new insurance. Renegotiation and contracting costs are often assumed to be proportional to the pure premium of risk that has to be insured. Therefore, if the contract is canceled at time $t = 1, 2, \ldots, T-1$, the insurer incurs a cost:

$$K(t) = cP(t,T) = cp(T-t),$$

for a given $c \geq 0$. The case where the insurer is allowed to cancel without any compensation to the insured corresponds to $c = 0$.

Finally, notice here that at time t, the estimated expected claim frequency, λ_t, as defined in (2), contains all the necessary information available in $\mathcal{H}_t$ to compute the posterior distribution of $\Lambda|\mathcal{H}_t$. Indeed, $\lambda_t = \lambda$ implies that $\Lambda|\mathcal{H}_t \sim Gamma(t_0 + t, \lambda(t_0 + t))$. Therefore, in the sequel we will denote $E[\cdot|\lambda_t = \lambda]$ for the expectation with respect to the posterior distribution at time t, given that $\lambda_t = \lambda$. In the next section, we derive the optimal policy with regard to canceling.

3 The canceling rule

At epoch $t = 1, 2, \ldots, T-1$, the insurer can decide whether he wants to continue the insurance. This implies that, at epoch t, he can use the information in $\mathcal{H}_t$ to choose between the following two options:

- Continue with the contract at least one more period. This implies that the surplus generated in $(t, t+1]$ equals

  $$C_{nc}(t) = p - S_{t+1} + S_t.$$

  After this period the surplus generated in $(t+1, T]$ depends on future decisions to cancel or continue with the contract in a certain period.

- Cancel the contract at epoch t. This implies that the surplus generated in $(t, t+1]$ is equal to the canceling cost, i.e.

$$C_c(t) = -cp(T-t),$$

and no surplus is generated in $(t+1, T]$.

Now let L denote the actual number of periods the contract sustains, i.e. the time until cancellation, or T otherwise. Then, the surplus generated by the contract over its lifetime equals:

$$\begin{aligned} Z_L &= \sum_{t=1}^{L-1} C_{nc}(t) + C_c(L) \\ &= pL - \sum_{n=1}^{N_L} X_n - cp(T-L). \end{aligned}$$

The aim is to determine the optimal L, given that canceling is costly, and that information becomes more reliable over time. A "naive" strategy for the insurer would be to cancel as soon as the updated claim frequency implies that the expected surplus that results from continuing with the contract until epoch T is lower than the surplus that results from canceling the contract, i.e. cancel as soon as:

$$\begin{aligned} & E\left[\sum_{i=t}^{T} C_{nc}(i) \,\middle|\, \lambda_t\right] < E[C_c(t) \mid \lambda_t] && \Leftrightarrow \\ & E\big[p(T-t) - S_T + S_t \mid \lambda_t\big] < -cp(T-t) && \Leftrightarrow \\ & p(T-t) - \lambda_t(T-t)\mu < -cp(T-t) && \Leftrightarrow \\ & \lambda_t > \frac{(1+c)p}{\mu}. \end{aligned}$$

In the sequel, we will denote:

$$\tilde{\lambda} := \frac{(1+c)p}{\mu} = (1+c)(1+\theta)\lambda_0. \tag{3}$$

Then the above strategy is of the control-limit type, i.e. cancel the contract as soon as the estimated claim frequency exceeds a certain, constant, limit $\tilde{\lambda}$. However, in early periods the information may still be relatively unreliable. Future estimates of the claim frequency may reveal that the contract is favorable. Whereas canceling is an irreversible decision, not canceling at the current period still leaves the option to cancel in later periods. The value of this option is not taken into account in the naive control-limit strategy.

In order to be able to take into account all future options, we use stochastic dynamic programming to determine the optimal policy with regard to canceling. The optimality equation that characterizes dynamic programming (see e.g. [8]) follows by considering the two options "canceling" and "not canceling". It is the option "not canceling" that yields the recursion forward in time. Therefore, we define the functions:

$$\begin{aligned} V_T(\lambda) &= 0, \\ V_t(\lambda) &= \max\left\{\, E[C_c(t)\,|\,\lambda_t = \lambda],\; E[C_{nc}(t) + V_{t+1}(\lambda_{t+1})\,|\,\lambda_t = \lambda]\,\right\}, \quad t = 0, \ldots, T-1, \end{aligned} \tag{4}$$

so that $V_t(\lambda)$ equals the expected surplus of the insurer generated in $[t, T]$, given that the optimal canceling policy is applied in $[t, T]$, that $\lambda_t = \lambda$, and that the contract was not canceled in $[0, t)$. The optimal expected surplus of the contract is then given by $V_0(\lambda_0)$.

In the sequel we will show that there exist $\tilde{\lambda}_t \geq \tilde{\lambda}$, for $t = 1, 2, \ldots, T-1$, such that it is optimal to cancel at epoch t iff $\lambda_t > \tilde{\lambda}_t$. We therefore first present the following proposition, which provides a recursive evaluation of the expected value of the surplus at time L, given that the optimal policy is applied.

Proposition 1. *When the optimal canceling policy is applied, the expected value of the insurer's surplus generated over the lifetime of the insurance, equals:*

$$V_0(\lambda_0) = -cpT + \mu \max\{\tilde{\lambda} - \lambda_0 + R_0(\lambda_0), 0\}, \tag{5}$$

where $\tilde{\lambda}$ is as defined in (3)*, and $R_0(.)$ is determined recursively as follows: For all $\lambda > 0$,*

$$\begin{aligned} R_{T-1}(\lambda) &= 0, \\ R_t(\lambda) &= E[\,\max\{\tilde{\lambda} - \lambda_{t+1} + R_{t+1}(\lambda_{t+1}), 0\}\,|\,\lambda_t = \lambda\,], \end{aligned} \tag{6}$$

for $t = T-2, T-3, \ldots, 0$, where, at time t, λ_{t+1} is the random variable $\frac{\lambda_t(t_0+t)+N_{t+1}-N_t}{t_0+t+1}$.

Proof. We show by induction that, for all $\lambda > 0$:

$$V_t(\lambda) = -cp(T-t) + \mu \max\{\tilde{\lambda} - \lambda + R_t(\lambda), 0\}, \qquad t = T-1, T-2, \ldots, 0. \tag{7}$$

At time $t = T-1$, there is only one period left, so that either $L = T-1$, or $L = T$. Given the expected surplus in the last period for these two options, this yields:

$$\begin{aligned}
V_{T-1}(\lambda) &= \max\big\{\ E[C_c(T-1)\,|\,\lambda_{T-1} = \lambda],\ E[C_{nc}(T-1)\,|\,\lambda_{T-1} = \lambda]\ \big\} \\
&= \max\Bigg\{\ -cp,\ E\Bigg[p - \sum_{i=N_{T-1}+1}^{N_T} X_i \,\Bigg|\, \lambda_{T-1} = \lambda\Bigg]\ \Bigg\} \\
&= \max\{-cp,\ p - \lambda\mu\} \\
&= p + \max\{-(1+c)p,\ -\lambda\mu\} \\
&= p - \mu\tilde{\lambda} + \mu \max\{\tilde{\lambda} - \lambda,\ 0\} \\
&= -cp + \mu \max\{\tilde{\lambda} - \lambda + R_{T-1}(\lambda),\ 0\}.
\end{aligned}$$

Now assume that equation (7) holds for $t+1, \ldots, T-1$. Applying (4) then yields:

$$\begin{aligned}
V_t(\lambda) &= \max\big\{\ E[C_c(t)\,|\,\lambda_t = \lambda],\ E[C_{nc}(t) + V_{t+1}(\lambda_{t+1})\,|\,\lambda_t = \lambda]\ \big\} \\
&= \max\Bigg\{-cp(T-t),\ E\Bigg[p - \sum_{i=N_t+1}^{N_{t+1}} X_i + V_{t+1}(\lambda_{t+1}) \,\Bigg|\, \lambda_t = \lambda\Bigg]\ \Bigg\} \\
&= \max\Big\{-cp(T-t),\ p - \lambda\mu - cp(T-t-1) \\
&\qquad + \mu E\big[\max\{\tilde{\lambda} - \lambda_{t+1} + R_{t+1}(\lambda_{t+1}), 0\}\,|\,\lambda_t = \lambda\big]\Big\} \\
&= -cp(T-t) + \max\Big\{0,\ (1+c)p - \lambda\mu + \mu R_t(\lambda)\Big\} \\
&= -cp(T-t) + \mu \max\{\tilde{\lambda} - \lambda + R_t(\lambda), 0\}.
\end{aligned}$$

Thus, the validity of (7) has been proven for $t = T-1, T-2, \ldots, 0$. $\square$

The above proposition yields the optimal expected surplus, i.e. the total expected surplus obtained when the optimal canceling policy is applied. From its proof, we can immediately infer the optimal canceling policy.

Corollary 1. *If the contract has not been canceled yet at time t, then the policy that maximizes the insurer's expected surplus in $(t, T]$ is as follows:*

- *cancel (resp. do not cancel) if* $\lambda_t > \tilde{\lambda} + R_t(\lambda_t)$ *(resp.* $\lambda_t < \tilde{\lambda} + R_t(\lambda_t)$*),*

- *either cancel or do not cancel if* $\lambda_t = \tilde{\lambda} + R_t(\lambda_t)$*.*

Proof. Suppose $\lambda_t = \lambda$. It follows immediately from the proof of Proposition 1 that:

- If $\tilde{\lambda} - \lambda + R_t(\lambda) < 0$ (resp. > 0), then the expected surplus that maximizes $V_t(\lambda)$ corresponds to the case of canceling (resp. not canceling) the contract at epoch t.

- If $\tilde{\lambda} - \lambda + R_t(\lambda) = 0$, then the two decisions yield the same expected surplus.

$\square$

From now on we assume that the insurer only cancels the contract if this decision yields an expected surplus that is strictly higher than the surplus corresponding to not canceling. In other words, in the case of indifference according to Corollary 1, the insurer is assumed not to cancel the contract. The following results can be easily adapted to the case where the insurer always cancels in the case of indifference. It is important, however, that this decision is taken consistently. It now follows immediately that.

Proposition 2. *The optimal contract length is given by:*

$$L = \begin{cases} t, & \text{if } \inf\{u \in \{0,1,\dots,T-1\} : \lambda_u - R_u(\lambda_u) > \tilde{\lambda}\} = t, \\ T, & \text{otherwise.} \end{cases} \tag{8}$$

In order to calculate the optimal canceling policy from (8), one has to solve the recursion in (6). As stated above, at time t, λ_{t+1} is a random variable. Therefore, in order to solve the recursion, we first determine the conditional distribution of λ_{t+1}, given λ_t.

Lemma 1. *For any* $\lambda > 0$*, and* $n \in \mathbb{N}$*, one has:*

$$P\left(\lambda_{t+1} = \frac{\lambda(t+t_0)+n}{t+t_0+1} \,\middle|\, \lambda_t = \lambda \right) = \binom{\lambda(t+t_0)+n-1}{n} \left(\frac{t+t_0}{t+t_0+1}\right)^{\lambda(t_0+t)} \left(\frac{1}{t+t_0+1}\right)^n.$$

Proof. The prior distribution of claim arrival rate Λ at epoch t, as given by (1), is characterized by the following two parameters:

$$\alpha(t) = t_0 + t,$$
$$\beta(t) = \lambda_0 t_0 + N_t.$$

Now let $\upsilon(.;\alpha(t),\beta(t))$ denote the pdf of this gamma distribution. Then for all $t = 0, 1, \ldots, T-1$, and $n \in \mathbb{N}$, one has:

$$\begin{aligned} P(N_{t+1} - N_t = n|\lambda_t = \lambda) &= \int_0^\infty P(N_{t+1} - N_t = n \,|\, \Lambda = x)\upsilon(x;\alpha(t),\beta(t))dx \\ &= \binom{\beta(t)+n-1}{n}\left(\frac{\alpha(t)}{\alpha(t)+1}\right)^{\beta(t)}\left(\frac{1}{\alpha(t)+1}\right)^n. \end{aligned} \tag{9}$$

Notice that it follows from (2) that, for given $\lambda_t = \lambda$, one has:

$$(\alpha(t),\beta(t)) = (t_0 + t, \lambda(t_0 + t)),$$
$$\lambda_{t+1} = \frac{\beta(t+1)}{\alpha(t+1)} = \frac{\beta(t) + N_{t+1} - N_t}{\alpha(t)+1}.$$

Now (9) immediately yields the desired result. □

In the sequel we show that, for all t, the function $R_t(.)$ is non-increasing. We therefore need the following lemma.

Lemma 2. *If $\lambda > \tau$, then $\lambda_{t+1} \,|\, \lambda_t = \lambda$ is stochastically larger than $\lambda_{t+1} \,|\, \lambda_t = \tau$.*

Proof. We have to show that:

$$P(\lambda_{t+1} > x|\lambda_t = \lambda) \geq P(\lambda_{t+1} > x|\lambda_t = \tau), \qquad \text{for all } x \in \mathbb{R}.$$

Define $\alpha = t_0 + t$, $\beta = \lambda\alpha$, and $\gamma = \tau\alpha$, and denote

$$z(n,\alpha,\beta) = \frac{\beta+n}{\alpha+1} \tag{10}$$
$$p(n,\alpha,\beta) = \binom{\beta+n-1}{n}\left(\frac{\alpha}{\alpha+1}\right)^{\beta}\left(\frac{1}{\alpha+1}\right)^n, \tag{11}$$

for $n \in \mathbb{N}$, $\alpha > 0$, and $\beta > 0$.

Let k be such that:

$$\{n \geq 0 : z(n,\alpha,\gamma) > x\} = \{k, k+1, \dots\}.$$

Since $\beta > \gamma$, it follows from (10) that:

$$\{n \geq 0 : z(n,\alpha,\beta) > x\} \supseteq \{k, k+1, \dots\}.$$

Hence, it is sufficient to show that

$$\sum_{n=k}^{\infty} p(n,\alpha,\beta) \geq \sum_{n=k}^{\infty} p(n,\alpha,\gamma), \qquad \text{for all } k \in \mathbb{N}. \tag{12}$$

First we prove that $p(n,\alpha,\beta) \geq p(n,\alpha,\gamma)$ implies that $p(n+1,\alpha,\beta) \geq p(n+1,\alpha,\gamma)$. Assume that $p(n,\alpha,\beta) \geq p(n,\alpha,\gamma)$. Then

$$\begin{aligned}
p(n+1,\alpha,\beta) &= \binom{\beta+n}{n+1}\left(\frac{\alpha}{\alpha+1}\right)^{\beta}\left(\frac{1}{\alpha+1}\right)^{n+1} \\
&= \frac{\beta+n}{(\alpha+1)(n+1)} p(n,\alpha,\beta) \\
&\geq \frac{\gamma+n}{(\alpha+1)(n+1)} p(n,\alpha,\gamma) \\
&= p(n+1,\alpha,\gamma).
\end{aligned}$$

Let $l = \min\{n \in \mathbb{N} : p(n,\alpha,\beta) \geq p(n,\alpha,\gamma)\}$. Then it follows that $p(n,\alpha,\beta) < p(n,\alpha,\gamma)$ for $0 \leq n < l$. It is clear that (12) holds for $k \geq l$, and for $0 \leq k < l$, we have

$$\sum_{n=k}^{\infty} p(n,\alpha,\beta) = 1 - \sum_{n=0}^{k-1} p(n,\alpha,\beta) \geq 1 - \sum_{n=0}^{k-1} p(n,\alpha,\gamma) = \sum_{n=k}^{\infty} p(n,\alpha,\gamma).$$

□

Proposition 3. *For all $t = 0, 1, \dots, T-1$, the function $R_t(.)$ is non-increasing on $\mathbb{R}_+$.*

Proof. Let $\lambda > \tau > 0$. For $t = T-1$, we have by definition that $R_{T-1}(\lambda) = R_{T-1}(\tau) = 0$. Now assume that the statement is true for $t+1, t+2, \dots, T-1$. It then follows that $\max\{\tilde{\lambda} - \lambda + R_{t+1}(\lambda), 0\}$ is an non-increasing function of λ. By Lemma 2, we know that $\lambda_{t+1}|\lambda_t = \lambda$ is stochastically larger than $\lambda_{t+1}|\lambda_t = \tau$. By a basic property on stochastic order relations (see e.g. [9], pp. 405) it follows immediately that $R_t(\lambda) \leq R_t(\tau)$. □

The fact that $R_t(.)$ is non-increasing allows to show that it is optimal for the insurer to cancel at a given epoch iff the number of observed claims exceeds a certain threshold value. These threshold values are presented in the following theorem.

Proposition 4. *It is optimal to cancel at epoch t if and only if $N_t > d_t$, where*

$$d_t = \max_{n\in\mathbb{N}} \left\{ \frac{\lambda_0 t_0 + n}{t_0 + t} \leq \tilde{\lambda} + R_t\left(\frac{\lambda_0 t_0 + n}{t_0 + t}\right) \right\}, \qquad \text{for } t = 1, \dots, T-1. \tag{13}$$

Moreover, these canceling limits satisfy $d_1 \leq d_2 \leq \cdots \leq d_{T-1}$.

Proof. We know from Corollary 1 that it is optimal not to cancel at time t iff:

$$\lambda_t \leq \tilde{\lambda} + R_t(\lambda_t). \tag{14}$$

Since $\lambda_t = \frac{\lambda_0 t_0 + N_t}{t_0 + t}$, and since $R_t(.)$ is non-increasing, it follows immediately that:

$$\lambda_t \leq \tilde{\lambda} + R_t(\lambda_t) \;\Leftrightarrow\; N_t \leq d_t, \qquad \text{for } t = 0, 1, \dots, T-1, \tag{15}$$

where d_t is given by (13).

It now remains to show that $d_t \leq d_{t+1}$, for $t = 1, \dots, T-2$. By definition, one has:

$$\frac{\lambda_0 t_0 + d_t}{t_0 + t} \leq \tilde{\lambda} + R_t\left(\frac{\lambda_0 t_0 + d_t}{t_0 + t}\right).$$

We now consider two cases.

(i) $\dfrac{\lambda_0 t_0 + d_t}{t_0 + t} \leq \tilde{\lambda}.$

In this case, it immediately follows that:

$$\frac{\lambda_0 t_0 + d_t}{t_0 + t + 1} \leq \frac{\lambda_0 t_0 + d_t}{t_0 + t} \leq \tilde{\lambda} \leq \tilde{\lambda} + R_{t+1}\left(\frac{\lambda_0 t_0 + d_t}{t_0 + t + 1}\right).$$

(ii) $\dfrac{\lambda_0 t_0 + d_t}{t_0 + t} > \tilde{\lambda}.$

This implies that $R_t\left(\frac{\lambda_0 t_0 + d_t}{t_0 + t}\right) > 0$. Since $\lambda_{t+1} \geq \frac{\lambda_0 t_0 + d_t}{t_0 + t + 1}$ if $\lambda_t = \frac{\lambda_0 t_0 + d_t}{t_0 + t}$, it follows from the definition of $R_t(\lambda)$, and the fact that $R_{t+1}(.)$ is non-increasing, that

$$R_t\left(\frac{\lambda_0 t_0 + d_t}{t_0 + t}\right) \leq \max\left\{ \tilde{\lambda} - \frac{\lambda_0 t_0 + d_t}{t_0 + t + 1} + R_{t+1}\left(\frac{\lambda_0 t_0 + d_t}{t_0 + t + 1}\right),\, 0 \right\}. \tag{16}$$

Since the left-hand side of (16) is strictly positive, it follows that:

$$\tilde{\lambda} - \frac{\lambda_0 t_0 + d_t}{t_0 + t + 1} + R_{t+1}\left(\frac{\lambda_0 t_0 + d_t}{t_0 + t + 1}\right) > 0.$$

Therefore, in both cases, it follows from (13) that $d_{t+1} \geq d_t$. $\square$

It now follows immediately that:

Corollary 2. *It is optimal to cancel at time t if and only if $\lambda_t > \tilde{\lambda}_t$, with $\tilde{\lambda}_t$ given by:*

$$\tilde{\lambda}_t = \frac{\lambda_0 t_0 + d_t}{t_0 + t}, \qquad t = 1, \ldots, T-1,$$

where d_t are as defined in (13).

It is seen from the above corollary, combined with (13) and (3), that in contrast to the naive policy, the canceling limit $\tilde{\lambda}_t$ not only depends on the premium density p, and the proportional canceling cost, but also on the time instance t, and the parameters λ_0 and t_0 of the prior distribution at date zero. Notice furthermore that, since $R_t(.) \geq 0$, Corollary 1 implies that:

$$\tilde{\lambda}_t \geq \tilde{\lambda}, \tag{17}$$

so that the naive policy always cancels earlier than the optimal policy.

In the remainder of this section we present some analytical results on the effect of the parameters on the optimal canceling policy, and illustrate these results in a numerical example.

Proposition 5. *Let $\mathcal{C}_1$ and $\mathcal{C}_2$ denote two insurance contracts with contract periods T_1 and $T_2 = T_1 - 1$, respectively* ($T_1 \in \mathbb{N},\ T_1 > 1$), *with all other contract specifications being equal for $\mathcal{C}_1$ and $\mathcal{C}_2$. Then*

$$R_t^{(1)}(\lambda) \geq R_t^{(2)}(\lambda), \qquad t = 0, 1, \ldots, T_2 - 1,\ \lambda > 0, \tag{18}$$

with

$$R_{T_i-1}^{(i)}(\lambda) = 0,$$
$$R_t^{(i)}(\lambda) = E\big[\max\{\tilde{\lambda} - \lambda_{t+1} + R_{t+1}^{(i)}(\lambda_{t+1}),\ 0\} \,\big|\, \lambda_t = \lambda\big],$$

for $\lambda > 0$, $i = 1, 2$, $t = 0, 1, \ldots, T_i - 2$, and with $\lambda_{t+1} = \frac{\lambda_t(t_0+t)+N_{t+1}-N_t}{t_0+t+1}$ and $\tilde{\lambda}$ defined by (3).

Proof. First we prove (18) for $t = T_2 - 1$.

$$\begin{aligned}
R^{(1)}_{T_2-1}(\lambda) &= E\big[\max\{\tilde{\lambda} - \lambda_{T_2} + R^{(1)}_{T_2}(\lambda_{T_2}), 0\} \,\big|\, \lambda_{T_2-1} = \lambda\big], \\
&= E\big[\max\{\tilde{\lambda} - \lambda_{T_1-1}, 0\} \,\big|\, \lambda_{T_1-2}\big] \\
&\geq 0 = R^{(2)}_{T_2-1}(\lambda).
\end{aligned}$$

Now assume that we have proven (18) for $t+1, t+2, \ldots, T_2 - 1$. This yields

$$\begin{aligned}
R^{(1)}_{t}(\lambda) &= E\big[\max\{\tilde{\lambda} - \lambda_{t+1} + R^{(1)}_{t+1}(\lambda_{t+1}), 0\} \,\big|\, \lambda_{t} = \lambda\big] \\
&\geq E\big[\max\{\tilde{\lambda} - \lambda_{t+1} + R^{(2)}_{t+1}(\lambda_{t+1}), 0\} \,\big|\, \lambda_{t} = \lambda\big] \\
&= R^{(2)}_{t}(\lambda).
\end{aligned}$$

Hence, (18) is valid for t and the result follows by induction. □

The above proposition immediately yields the following result.

Corollary 3. *Let $d^{(1)}_t$ and $d^{(2)}_t$ denote the optimal canceling limits for the contracts $\mathcal{C}_1$ and $\mathcal{C}_2$ mentioned in Proposition 5 respectively. Then*

$$d^{(1)}_t \geq d^{(2)}_t, \qquad t = 1, \ldots, T_2 - 1.$$

It is clear from (6) and Corollary 1 that the optimal policy only depends on θ and c through $\tilde{\lambda} = (1+c)(1+\theta)\lambda_0$. The following proposition therefore provides the effect of an increase in $\tilde{\lambda}$ on $R_t(.)$.

Proposition 6. *Let $\tilde{\lambda}^{(1)} > \tilde{\lambda}^{(2)}$ and define*

$$\begin{aligned}
R^{(i)}_{T-1}(\lambda) &= 0, \\
R^{(i)}_{t}(\lambda) &= E\big[\max\{\tilde{\lambda}^{(i)} - \lambda_{t+1} + R^{(i)}_{t+1}(\lambda_{t+1}), 0\} \,\big|\, \lambda_{t} = \lambda\big],
\end{aligned}$$

for $\lambda > 0$, $t = 1, 2, \ldots, T-2$, $i = 1, 2$, and with $\lambda_{t+1} = \frac{\lambda_t(t_0+t)+N_{t+1}-N_t}{t_0+t+1}$. Then

$$R^{(1)}_t(\lambda) \geq R^{(2)}_t(\lambda), \qquad t = 1, 2, \ldots, T-1,\ \lambda > 0. \tag{19}$$

Proof. By definition inequality (19) is valid for $t = T - 2$. Now assume (19) is valid for $t = \tau + 1$. Then, for $\lambda > 0$, we have

$$\begin{aligned} R_\tau^{(1)}(\lambda) &= E\big[\max\{\tilde{\lambda}^{(1)} - \lambda_{\tau+1} + R_{\tau+1}^{(1)}(\lambda_{\tau+1}), 0\} \,\big|\, \lambda_\tau = \lambda\big] \\ &\geq E\big[\max\{\tilde{\lambda}^{(2)} - \lambda_{\tau+1} + R_{\tau+1}^{(2)}(\lambda_{\tau+1}), 0\} \,\big|\, \lambda_\tau = \lambda\big] \\ &= R_\tau^{(2)}(\tau). \end{aligned}$$

Hence, (19) is valid for $t = \tau$, and, consequently, the result follows by induction. □

Corollary 4. *Consider two insurance contracts $\mathcal{C}_1$ and $\mathcal{C}_2$ that have the same contract length T and prior distribution for Λ. Let $\tilde{\lambda}^{(1)}$ and $\tilde{\lambda}^{(2)}$ denote the corresponding control-limits of the naive strategy as defined by* (3)*, and let $d_t^{(1)}$ and $d_t^{(2)}$ denote the optimal canceling limits of contracts $\mathcal{C}_1$ and $\mathcal{C}_2$ that correspond to* (13)*, respectively. If $\tilde{\lambda}^{(1)} > \tilde{\lambda}^{(2)}$, then*

$$d_t^{(1)} \geq d_t^{(2)}, \qquad t = 1, \ldots, T - 1.$$

We conclude this section with a numerical example of the effect of the different parameters on the canceling limits.

Example

Consider the five insurance contracts with contract parameters shown in Table 1. Contracts $\mathcal{C}_2$–$\mathcal{C}_5$ differ from $\mathcal{C}_1$ only in one single parameter value. Figure 1 shows $R_t(\lambda) + \tilde{\lambda}$ ($t = 1, 2, \ldots, 8$) for contract $\mathcal{C}_1$ as a function of λ for contract $\mathcal{C}_1$. It follows from Corollary 1 that the intersections with λ, the dotted line, yield the values of λ at which the insurer is indifferent between canceling and not canceling at epoch $t = 1, 2, \ldots, 8$. These intersections can be translated into canceling limits d_t using formula (13). The canceling limits of contracts $\mathcal{C}_1$–$\mathcal{C}_5$ are shown in Table 2.

The second column marked by $\mathcal{C}_1^*$ presents the canceling limits that correspond to the naive canceling rule applied to contract $\mathcal{C}_1$. Notice that the naive canceling rule cancels earlier than the optimal canceling rule (see (17)). The canceling limits of the naive and optimal canceling rule for $\mathcal{C}_1$ are equal only for $t = T - 1 = 9$, since then one has $R_t(\lambda) = 0$.

The canceling limits of contracts $\mathcal{C}_2$ and $\mathcal{C}_3$ are both higher or equal to the limits of $\mathcal{C}_1$. The reason for this is that both an increase of c or an increase of θ result in an increase

Parameter	Contract				
	$\mathcal{C}_1$	$\mathcal{C}_2$	$\mathcal{C}_3$	$\mathcal{C}_4$	$\mathcal{C}_5$
T	10	10	10	10	15
λ_0	2	2	2	2	2
t_0	8	8	8	10	8
c	0.03	0.06	0.03	0.03	0.03
θ	0.10	0.10	0.15	0.10	0.10

Table 1: Contract parameters

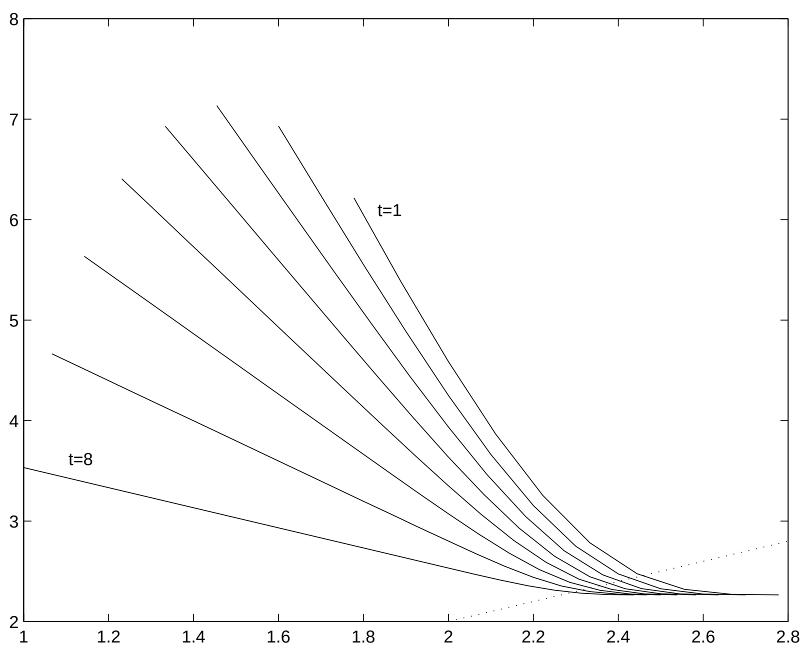


Figure 1: $\tilde{\lambda} + R_t(\lambda)$ for $\mathcal{C}_1$ as a function of λ.

t	d_t					
	$\mathcal{C}_1$	$\mathcal{C}_1^*$	$\mathcal{C}_2$	$\mathcal{C}_3$	$\mathcal{C}_4$	$\mathcal{C}_5$
1	8	5	8	9	8	8
2	10	7	10	11	10	10
3	12	9	13	13	12	13
4	14	12	15	15	15	15
5	16	14	17	17	17	17
6	18	16	19	20	19	19
7	20	18	21	22	21	21
8	22	21	23	24	23	23
9	23	23	24	25	24	25
10	-	-	-	-	-	27
11	-	-	-	-	-	29
12	-	-	-	-	-	31
13	-	-	-	-	-	33
14	-	-	-	-	-	34

Table 2: Canceling limits

of $\tilde{\lambda}$, while (λ_0, t_0) remains unchanged, which leads to higher canceling limits according to Proposition 6.

The value of t_0 in contract $\mathcal{C}_4$ is higher ($t_0 = 10$) than the value of t_0 in contract $\mathcal{C}_1$ ($t_0 = 8$). Consequently, the posterior distribution in the case of contract $\mathcal{C}_4$ is updated more conservatively, and hence more "evidence", i.e. a higher canceling limit, is needed to convince the insurer that it is optimal to cancel the contract.

Finally, contract $\mathcal{C}_5$ has a longer contract period than contract $\mathcal{C}_1$. This implies that the canceling limits for contract $\mathcal{C}_5$ are higher (see Proposition 5). The intuition for this difference is that, since the remaining contract length is longer, both the canceling costs and the potential loss when canceling an actually favorable contract are higher.

4 Implications for Premium Setting

For any given safety loading θ, proportional canceling cost c, and maximal contract length T, the canceling limits $d(T) := (d_1, \ldots, d_{T-1})$, as defined in (13), can be determined. This implies that in case of large, client tailored, insurance, the contract can specify a combination of θ, c, T, and $d(T)$. Here, for example, a higher θ or a higher c will allow for higher canceling limits d_t, and vice versa. Obviously, these contract specifications affect the expected value of the insurer's net profit, as well as the insured's expected utility of the contract. Therefore, in the sequel, we study how the contract specifications affect the relation between expected costs and expected premium income.

In the standard classical risk process, the expected value of the insurer's surplus at a given time t is proportional to the expected value of the total claim height up to time t. More precisely, in the absence of canceling options, the surplus generated by the contract in $[0, t]$ equals:

$$Z_t = pt - \sum_{i=1}^{N_t} X_i,$$

so that the expected surplus at the end of the planning period equals:

$$\begin{aligned} E[Z_T] &= (1+\theta)\lambda_0 \mu T - \lambda_0 \mu T \\ &= \theta E[S_T]. \end{aligned}$$

When the option to cancel is available, the above described relationship no longer holds. In general, the effective safety loading θ_L—which equals the fraction of expected surplus over expected costs, taking into account the optimal canceling strategy and the canceling costs—differs from the safety loading θ used to determine the premium density at date zero.

We therefore now study the effect of the canceling option on the effective safety loading. As before, we denote L for the random variable that denotes the optimal lifetime of the contract, i.e. the time at which it is canceled, or the time until it expires, T, if it is not canceled. The insurer's optimal canceling policy implies that:

$$L = \begin{cases} t, & \text{if } \inf\{u \in \mathbb{N} : N_u > d_u\} = t < T, \\ T, & \text{if } N_u \leq d_u \text{ for all } u \leq T-1. \end{cases}$$

Now, for any given canceling rule $d(T) = (d_1, \ldots, d_{T-1})$, and (subjective) probability distribution of Λ, the distribution of L can be determined. This distribution can then be used to determine the net expected profit generated by the contract, taking into account its actual lifetime and the cancellation costs. Since in this section, all expectations are taken at date zero, we will omit the notation $|\mathcal{H}_0$, and denote $E[.]$ for the expected value with respect to the subjective distribution used at date zero.

Since L is a stopping time w.r.t. the process $\{N_t : t \geq 0\}$ it follows that $E[N_L] = E[\Lambda L]$, and consequently the expected value of the insurer's surplus at time L, equals:

$$\begin{aligned} E[Z_L] &= E\Bigg[pL - \sum_{i=1}^{N_L} X_i - cp(T-L) \Bigg], \\ &= pE[L] - E[\Lambda L]\mu - cp(T - E[L]). \end{aligned}$$

Taking into account the canceling costs, and the actual lifetime of the contract, the *effective safety loading* for the contract equals:

$$\begin{aligned} \theta_L &= \frac{E[Z_L]}{E[S_L] + cp(T - E[L])}, \\ &= \frac{pE[L]}{E[\Lambda L]\mu + cp(T - E[L])} - 1. \end{aligned}$$

Notice here that, even when canceling costs are zero, i.e. $c = 0$, the effective safety loading in general still deviates from the safety loading θ. Indeed, we then have that:

$$\theta_L = \frac{(1+\theta)\lambda_0 \mu E[L]}{E[\Lambda L]\mu} - 1,$$

which only equals θ if $E[\Lambda L] = \lambda_0 E[L] = E[\Lambda]E[L]$. It is clear that, since L depends on Λ, this equality need not be satisfied.

To conclude, we present a numerical example to illustrate the effect of the availability of the option to cancel on: the lifetime distribution, the insurer's expected surplus at the end of the contract, and the effective safety loading.

Example

Without any loss of generality we will assume that $\mu = 1$, hence the premium density is given by $p = (1+\theta)E[\Lambda] = (1+\theta)\lambda_0$.

We consider an insurance contract with a fixed time horizon of five periods ($T = 5$) and the option, available to the insurer, to cancel at the end of each period at canceling costs equal to 3% ($c = 0.03$) of the premium for the remaining contract period. We assume that the insurer uses a $Gamma(4, 16)$ ($\lambda_0 = 4, t_0 = 4$) prior distribution for the Poisson claim arrival parameter Λ, hence $E[\Lambda] = 4$ and $V[\Lambda] = 1$.

Next, we consider eight different premiums resulting from safety loadings:

$$\theta = \theta_n := 0.025 \times n, \qquad \text{for } n = 1, 2, \ldots, 8.$$

Figure 2(a) presents the canceling limits $d_t, t = 1, \ldots, 4$, for these eight different values of θ. We see that, for each t, the canceling limit d_t increases as θ increases. This also immediately follows from Corollary 4, since an increase in θ implies an increase in $\tilde{\lambda}$. In words, this result can be explained by two effects:

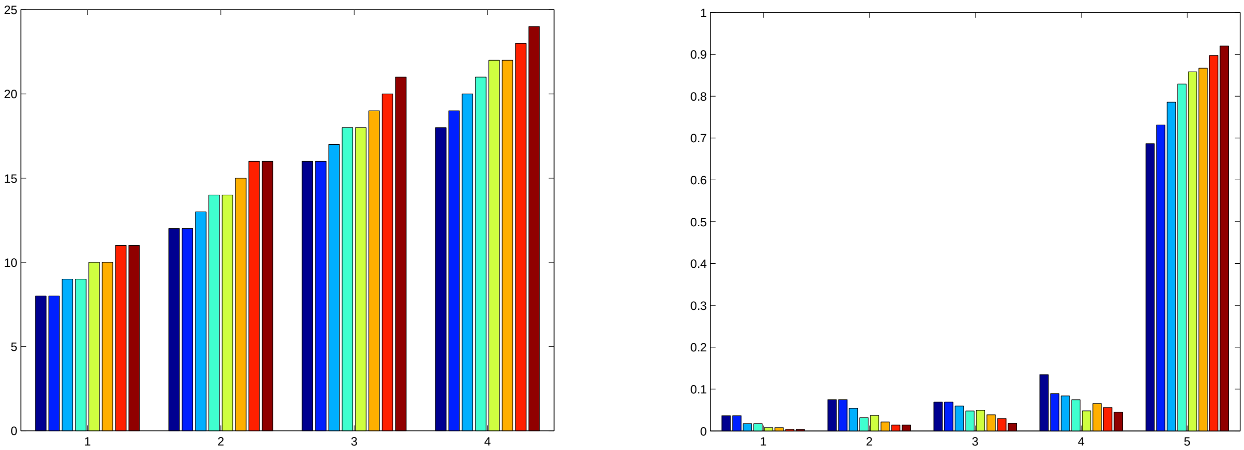
- A higher θ implies a higher premium density p. Since θ does not affect the claim height, the expected surplus resulting from not canceling is affected positively.

- A higher θ implies higher canceling costs, so that the surplus resulting from canceling is affected negatively.

The canceling limits for these different values of θ can be used to determine the probability that the contract will sustain for T periods (see Figure 2(b)), as well as the expected lifetime. Both are increasing functions of θ.

Table 3 shows some important statistics of the insurance contract for the insurer. The expected number of claims that the insurer has to cover is given by $E[\Lambda L]$, which is not equal to $E[\Lambda]E[L]$ since, obviously, the lifetime distribution depends on the claim arrival rate Λ. In our example we see that $E[\Lambda L] < E[\Lambda]E[L]$. This can be explained by the fact that if the *actual* claim arrival rate λ is high, relative to the prior expectation λ_0, then the contract is likely to be canceled before the expiration date.

Notice that the effective safety loading θ_L differs significantly from the safety loading used to determine the premium density at date zero. The overall effect of an increasing safety loading θ on the insurer's expected surplus, $E[Z_L]$, is positive, as could be expected.



(a) Canceling limits d_t.

(b) Lifetime distribution: $P(L = t)$.

Figure 2: Canceling limits and lifetime distribution for eight different premium levels. The bars in both figures correspond to safety loadings $\theta_1, \ldots, \theta_8$, from the left to the right.

θ	$E[L]$	$E[\Lambda L]$	$E[Z_L]$	θ_L
0.025	4.3598	16.8375	0.9591	0.0567
0.050	4.4046	17.0335	1.3910	0.0813
0.075	4.5664	17.7855	1.7940	0.1006
0.100	4.6666	18.2621	2.2268	0.1216
0.125	4.7116	18.4794	2.6838	0.1449
0.150	4.7621	18.7303	3.1425	0.1675
0.175	4.8294	19.0693	3.6049	0.1888
0.200	4.8639	19.2447	4.0822	0.2119

Table 3: Statistics for the contracts with different levels of θ.

5 Summary

In this paper we have analyzed a risk process where the insurer has the option to stop offering the insurance before the end of the planning horizon. Such a type of risk process is especially suitable to model the situation where the insurer does not have enough knowledge about the risk process to be able to accurately predict the risks involved. In particular for large industrial risks, where insurance is client specific, both insurer and insured can benefit from longer term contracts since it saves renegotiation and contracting costs. However, if the contract covers multiple periods, then misjudging these risks can imply that the insurer is committed to an unfavorable contract for a lengthy period. The option to cancel the contract is a very valuable option in such a case. For the insured, however, this option to cancel generates an additional risk, since there is no certainty that all future claims will be covered by the insurer. To compensate for this inconvenience canceling costs have to be paid to the insured by the insurer.

Also in case of introduction of a new insurance product, offered to a large group of insureds, information at the start of the planning period may be relatively unreliable, so that, after a number of time periods the insurer may want to stop offering the insurance at the given premium system.

In both cases, the insurer is faced with the decision problem whether to cancel the contract or continue with it. This decision is affected by the history of the risk process so far, since this information enables the insurer to more accurately predict the future of the risk process.

For the case of a Poisson claim arrival process with unknown parameter and prior gamma distribution, we have derived a solution method, using stochastic dynamic programming, resulting in the optimal canceling policy. The resulting canceling rule depends on the history of the process only through the total number of claims up to that point in time. We found that, at a certain decision point, the number of remaining periods is significant for the insurer's decision whether to cancel or not. More specifically, we showed that if the canceling rule prescribes to cancel the contract if more than one periods are remaining, then it will also prescribe to cancel if there is one period remaining, given that the history of the risk process is identical in the two cases.

Taking a certain canceling policy and the insurer's prior distribution for the claim arrival rate as a starting-point, contract specifications such as the premium and canceling costs provide an instrument for the insurer to influence the canceling probability and the expected profit. In case of precommitment of the insurer, the insured on the other hand, has to evaluate the expected utility of a contract offered by the insurer, including the canceling options, and decide whether it is favorable compared to having no insurance contract.

References

[1] Berg, M. (1987). Reliability control for mission-critical items. *Naval Research Logistics* **34,** 417–429.

[2] Bowers, N., Gerber, H., Hickmann, J., Jones, D. and Nesbitt, C. (1986). *Actuarial Mathematics.* The Society of Actuaries.

[3] Bühlmann, H. (1970). *Mathematical Methods in Risk Theory.* Springer-Verlag.

[4] De Vylder, F. (1996). *Advanced Risk Theory: a self-contained introduction.* Editions de l'Université de Bruxelles, Brussels.

[5] Gerber, H. (1979). *An Introduction to Mathematical Risk Theory.* S.S. Huebner Foundation Monograph Series. Richard D. Irwin, Inc.

[6] Goovaerts, M., De Vylder, F. and Haezendonck, J. (1984). *Insurance Premiums.* North Holland, Amsterdam.

[7] Grandell, J. (1991). *Aspects of Risk Theory.* Springer-Verlag.

[8] Ross, S. (1983). *Introduction to Stochastic Programming.* Academic Press.

[9] Ross, S. (1996). *Stochastic Processes* 2nd ed. John Wiley & Sons.